Prostate cancer screening with prostate-specific antigen (PSA) test: a systematic review and meta-analysis

Dragan Ilic,1 Mia Djulbegovic,2 Jae Hung Jung,3 Eu Chang Hwang,4,5 Qi Zhou,6 Anne Cleves,7 Thomas Agoritsas,6,8 Philipp Dahm5

ABSTRACT

OBJECTIVE
To investigate the efficacy and safety of prostate-specific antigen (PSA) testing to screen for prostate cancer.

DESIGN
Systematic review and meta-analysis.

DATA SOURCES
Electronic search of Cochrane Central Register of Controlled Trials, Web of Science, Embase, Scopus, OpenGrey, LILACS, and Medline, and search of scientific meeting abstracts and trial registers to April 2018.

ELIGIBILITY CRITERIA FOR SELECTING STUDIES
Randomised controlled trials comparing PSA screening with usual care in men without a diagnosis of prostate cancer.

DATA EXTRACTION
At least two reviewers screened studies, extracted data, and assessed the quality of eligible studies. A parallel guideline committee (BMJ Rapid Recommendation) provided input on the design and interpretation of the systematic review, including selection of outcomes important to patients. We used a random effects model to obtain pooled incidence rate ratios (IRR) and, when feasible, conducted subgroup analyses (defined a priori) based on age, frequency of screening, family history, ethnicity, and socioeconomic level, as well as a sensitivity analysis based on the risk of bias. The quality of the evidence was assessed with the GRADE approach.

RESULTS
Five randomised controlled trials, enrolling 721 718 men, were included. Studies varied with respect to screening frequency and intervals, PSA thresholds for biopsy, and risk of bias. When considering the whole body of evidence, screening probably has no effect on all-cause mortality (IRR 0.99, 95% CI 0.98 to 1.01; moderate certainty) and may have no effect on prostate-specific mortality (IRR 0.96, 0.85 to 1.08; low certainty). Sensitivity analysis of studies at lower risk of bias (n=1) also demonstrates that screening seems to have no effect on all-cause mortality (IRR 1.0, 0.98 to 1.02; moderate certainty) but may have a small effect on prostate-specific mortality (IRR 0.79, 0.69 to 0.91; moderate certainty). This corresponds to one less death from prostate cancer per 1000 men screened over 10 years. Direct comparative data on biopsy and treatment related complications from the included trials were limited. Using modelling, we estimated that for every 1000 men screened, approximately 1, 3, and 25 more men would be hospitalised for sepsis, require pads for urinary incontinence, and report erectile dysfunction, respectively.

CONCLUSIONS
At best, screening for prostate cancer leads to a small reduction in disease-specific mortality over 10 years but has does not affect overall mortality. Clinicians and patients considering PSA based screening need to weigh these benefits against the potential short and long term harms of screening, including complications from biopsies and subsequent treatment, as well as the risk of overdiagnosis and overtreatment.

SYSTEMATIC REVIEW REGISTRATION
PROSPERO registration number CRD42016042347.

Introduction
Prostate cancer is the second most common cancer and the fifth leading cause of cancer-associated mortality among men worldwide.1 Screening for prostate cancer with serum prostate-specific antigen (PSA) aims to detect prostate cancer at an early, interventional stage amenable to curative treatment and reduction in overall and disease-specific mortality.2 3 However, the evidence has so far not demonstrated that screening for prostate cancer saves lives.4 5 Instead, screening may be associated with increased harms such as overdiagnosis and complications of treatment for indolent disease.2 5 Nevertheless, screening for prostate cancer remains highly controversial because of limitations in randomised trials including contamination and under-representation of black men. Difficulty of shared, informed decision-making between patients and primary care providers about PSA screening may also contribute to practice variations.6 7 Recently, the US Preventive Services Task Force

WHAT IS ALREADY KNOWN ON THE SUBJECT
Prostate-specific antigen (PSA) screening leads to increased prostate cancer diagnoses. Although it is recommended by some clinical practice guidelines, PSA screening remains controversial. It is unclear whether screening improves overall and disease-specific mortality, the most critical outcomes for patients, or whether the overall benefits of screening outweigh the potential harms and costs of overdiagnosis and overtreatment. The recent publication of a large cluster-randomised controlled trial provides new evidence.

WHAT THIS STUDY ADDS
At best, screening for prostate cancer may result in a small absolute benefit in disease-specific mortality over 10 years but does not improve overall mortality. These benefits need to be weighed against the potential short and long term harms of PSA screening, including complications from biopsies and subsequent treatment and the risk of overdiagnosis and overtreatment.
Box 1: Linked articles in this BMJ Rapid Recommendation cluster

- Summary of the results from the Rapid Recommendation process
- Systematic review of the values and preference of men considering PSA screening
- MAGICapp (https://app.magicapp.org/public/guideline/n32gkL)
- Expanded version of the results with multilayered recommendations, evidence summaries, and decision aids for use on all devices

(USPSTF) updated their recommendation statement, changing it from a grade D (recommendation against PSA based screening for prostate cancer) to a grade C recommendation (advocating for an individualised approach to screening).1,8,9

Our prior systematic review and meta-analysis that evaluated the effectiveness of prostate cancer screening included five studies that enrolled 341 342 patients.5 In this 2013 Cochrane review we demonstrated that PSA screening led to an increase in prostate cancer diagnoses but did not reduce overall nor disease-specific mortality. Since the publication of this review, several of the included studies have reported outcomes at extended follow-up. In addition, a new, large trial has been published.10 Therefore, the effectiveness of prostate cancer screening based on the current best evidence is uncertain. In contrast to the recently published USPSTF evidence report, we included and analysed all relevant screening trials using data from the longest available follow-up.9

This systematic review is part of the BMJ Rapid Recommendations project, a collaborative effort from the MAGIC research and innovation programme (www.magicproject.org) and The BMJ. The aim of the project is to respond to new potentially practice changing evidence and provide a trustworthy practice guideline in a timely manner.11 In our case, the stimulus was the recent Cluster Randomised Trial of PSA Testing for Prostate Cancer (CAP),10 which randomised over 419 357 men at 573 primary care practices in the United Kingdom to PSA screening versus usual management.10 In light of this new evidence, we conducted an update of prior systematic reviews by our group to address the potential benefits and harms of PSA based screening.3,5 This systematic review informed the parallel guideline published in a multilayered electronic format on bmj.com and MAGICapp (see box 1).

**Methods**

**Protocol registration**

The protocol for this systematic review was registered with PROSPERO (CRD42016042347).16

**BMJ Rapid Recommendation and patient involvement**

In accordance with the BMJ Rapid Recommendations process,11 a guideline panel provided critical input and guidance during the review process, which included identifying populations, subgroups, and outcomes of interest. The panel consisted of general practitioners, urologists, methodologists, and men eligible for screening. These eligible men received personal training and support to optimise contributions throughout the guideline development process. They were full members of the guideline panel and contributed to the selection and prioritisation of outcomes. They also contributed to the assessment of values and preferences and provided critical feedback to the systematic review protocol and BMJ Rapid Recommendations manuscript.12,13

**Search strategy**

A trained medical librarian performed electronic searches of the Cochrane Central Register of Controlled Trials (via Wiley), Web of Science, Embase, Scopus, OpenGrey, LILACS, Medline (via Ovid) and PubMed (via National Library of Medicine) databases from their inception through to April 2018 (see appendix 1). Additionally, we scanned the reference lists of published narrative and systematic reviews to identify any potential studies not retrieved by our electronic search. In an effort to find unpublished studies, we also hand searched abstracts from the annual meetings of American Urological Association, American Society of Clinical Oncology, and European Urological Association from 2013 (the latest search date of our most recent systematic review5) through to April 2018. To identify ongoing trials, we used the International Clinical Trials Registry Platform (ICTRP) and ClinicalTrials.gov search portals.

The randomised controlled trials reported limited data on the harms of screening compared with no screening. We therefore searched for follow-up evidence from the intervention arms of included trials and follow-up publications. We used the Finnish arm of the European Randomised Study of Screening for Prostate Cancer (ERSPC) for quality of life data14 and...
false positive rates. We estimated false negative rates among men with a low PSA concentration from a follow-up cohort study of the Prostate Cancer Prevention Trial.17 We extracted complication rates from prostate cancer treatment modalities from the Prostate Testing for Cancer and Treatment (ProtecT) trial, which enrolled patients from the CAP trial who were diagnosed with prostate cancer and randomised to active monitoring, radical prostatectomy, or radical radiotherapy with hormones. Similarly, we obtained complication rates from biopsies from the Prostate Biopsy Effects cohort study nested within the ProtecT trial. By modelling the likelihood of prostate biopsy and treatment related complications among men who underwent screening compared with those who did not.

Study selection
After removing duplicates, we imported all citations into Covidence (www.covidence.org). Two study team members independently reviewed each citation and abstract according to a priori selection criteria. For studies that advanced beyond this stage, two study team members then performed independent full-text reviews. To determine eligibility, studies in non-English language were translated. References reporting on the same trial were mapped to unique study identifiers.

We considered all randomised controlled trials in any language comparing PSA based screening with usual care in men without a diagnosis of prostate cancer. Outcomes of interest were all-cause mortality, disease-specific mortality, prostate cancer incidence and stage (categorised as localised and advanced cancer), quality of life, false positive and false negative results, and harms related to screening.

Data extraction and risk of bias assessment
For each study selected for inclusion, two team members independently extracted data using a previously tested standardised form. Data collected were (a) general study information (authors, study year, publication type, country of origin, enrolment period, inclusion and exclusion criteria, and funding source); (b) study population details (ethnicity, sample size, age, and duration of follow-up); (c) details of the intervention and comparator (PSA, digital rectal examination, frequency of screening, and definition of control treatment or usual care); and (d) outcomes of interest as listed above. Inconsistencies in extracted data were resolved between the two team members through discussion, with a third team member serving as arbitrator. For the ERSPC study, we focused on the “core” group, which best corresponded to the target population of the accompanying Rapid Recommendation. The ERSPC study recruited men aged 50–74 years, but focused its data analysis on men in the core age group of 55–69 years.

To assess the risk of bias of the selected studies, we used Cochrane’s risk of bias tool on an outcome-specific basis. Two team members independently evaluated each randomised controlled trial across several domains by relying on the information presented in the study report, available protocols, or secondary publications. No assumptions were made. These domains included sequence generation and concealment of allocation (selection bias), blinding of participants and personnel (performance bias), blinding of outcome assessors (detection bias), and completeness of follow-up (attrition bias). In addition, we assessed rates of adherence to the screening intervention and rates of contamination of the usual care arm. For each domain, two individual team members judged whether the risk of bias was low, high, or unclear. Any disagreements were reconciled by a third team member.

We rated the confidence in the estimates of effect for each outcome according to the GRADE (Grading of Recommendations, Assessment, Development, and Evaluation) approach, taking into account study limitations (risk of bias), inconsistency, imprecision, indirectness, and publication bias. For each comparison, two team members independently rated the certainty of the effect estimates (that is, quality of evidence) for each outcome as high, moderate, low, or very low. We resolved discrepancies by consensus and, if needed, by arbitration by a third team member. The GRADE Summary of Findings table was generated using the MAGICapp platform (www.magicapp.org). The trials varied in their duration of follow-up from 10 to 20 years, and we pooled relative estimates of effect at the longest available follow-up time. To generate absolute effect size estimates, we consistently used the baseline risk from the CAP trial as it provided the most contemporary and therefore applicable control event rate in the absence of screening.10 We calculated absolute effect size estimates at 10 years, which we determined to be an appropriate time horizon to base individuals’ and panel members’ deliberations on. We also explored how the absolute effect estimates varied according to baseline risks at different time horizons using data from the ERSPC trial.

Data synthesis and primary analysis
When person-years data were available, we used incidence rate ratios with 95% confidence intervals to express dichotomous outcomes; alternatively, we used risk ratios. When needed, we approximated risk ratio to incidence rate ratio if the event rate was low (<10%) and the sample size was large; in the case that the event rate was a little over 10% (~12% prostate cancer incidence in Stockholm study23), we conducted a sensitivity analysis excluding this study. To account for different variances of the treatment effect due to clustering in the CAP trial,10 we used the adjusted incidence rate ratio from the generalised regression approach for the outcome measure.

We conducted meta-analyses and pooled the effect estimates using DerSimonian and Laird’s inverse of
variance random effects model and presented the results in forest plots.26 Following GRADE guidance, statistical heterogeneity was determined using the Q statistic and $I^2$ which was interpreted as follows; (a) 0-40% may not be important, (b) 30-60% may indicate moderate heterogeneity, (c) 50-90% indicates substantial heterogeneity, and (d) 75-100% indicates considerable heterogeneity.25 Decisions about downgrading for inconsistency were based on clinical relevance according to the clinical practice guideline that this systematic review supports.26, 27 Regardless of the observed statistical heterogeneity, and when the evidence was available, we conducted the following pre-specified subgroup analyses: age (50-54, 55-59, 60-64, 65-69, and ≥70 years), frequency of screening (periodic vs one-time), family history (present vs absent), ethnicity (African descent vs not), and socioeconomic level (higher vs lower) using lower levels of education (primary education only) as a proxy of lower socioeconomic status.28 For subgroup analyses, we used the random effects meta-regression approach to test for interaction. We also conducted a sensitivity analysis limited to studies assessed as having a lower risk of bias. We planned to use funnel plots to examine the extent of publication bias for outcomes if there were 10 or more studies included, which was not the case.14 Data were analysed using SAS (version 9.4, Cary, NC, US) and STATA software (Version 14.2, Texas, US). For all aspects of the review, we followed the guidelines of the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) checklist.29

Results
Description of included studies
Our electronic search identified 10,982 references. Of these, 5,385 were excluded as duplicates, leaving 5,597 for screening. We excluded 5,488 during the initial screening phase based on the title and abstract. For the remaining 109 studies, we undertook full-text screening and eliminated 77 studies for reasons including (a) interventions or comparators were not aligned with screening versus no screening (n=28), (b) not randomised controlled trials (n=21), (c) duplicate studies (n=12), (d) secondary analyses of previously reported data (n=10), (e) study participants did not meet eligibility criteria (n=4), and (f) outcomes were not aligned with primary or secondary outcomes of interest (n=2). We excluded the Norköping trial based on methodological grounds as it was quasi-randomised.30 This left 32 references which mapped to five unique studies that ultimately met our inclusion criteria and were included in the meta-analysis. Figure 1 describes the study selection process.

These five randomised controlled trials enrolled a total of 7,2178 participants. The CAP study was conducted in the United Kingdom, recruiting 573 primary care practices.10 The remaining four trials included the ERSPC study,31 the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening trial conducted in the US,32 and trials conducted in Canada (Quebec)33 and Sweden (Stockholm).34 The ERSPC study was a multicentre study across eight European countries.31

Among the included studies, the age of men enrolled ranged from 40 to 80 years. Screening methods included PSA alone and PSA combined with digital rectal examination. PSA thresholds to indicate further investigation via biopsy differed across studies, as did the screening interval (varying from one-time, every two years or more, to annual screening). Table 1 and appendix 2 provide additional details of the included studies.

All included studies were assessed as being potentially susceptible to performance bias because of the nature of the intervention. The CAP study had flawed allocation concealment (general practices consented after randomisation and group assignment) and adherence (only 40% of those randomised to screening adhered to it).10 The PLCO study had poor allocation concealment and major contamination.32 The two other smaller studies had further issues across several domains.23, 33 Figure 2 summarises the risk of bias assessment for each study. The summary of findings for the entire body of evidence is detailed in table 2. Of the five included studies, only the ERSPC study was found to have a lower risk of bias than the rest of the evidence from other trials.31 We therefore present a sensitivity analysis based on selected evidence from ERSPC in table 3.

**Main analysis including all eligible trials**

**Effect of PSA screening on all-cause mortality**

Based on four randomised controlled trials reporting this outcome, screening probably has no effect on all-cause mortality (incidence rate ratio (IRR) 0.99 (95% CI 0.98 to 1.01), $I^2=0\%$, moderate quality evidence; fig 3, table 2).10 23 31 32 This corresponds to one less death from any cause (95% CI 3 fewer to 1 more) per 1000 participants screened. We downgraded the quality of evidence for risk of bias.

---

![Flow diagram of study selection for systematic review](https://www.bmj.com/)

**Fig 1** Flow diagram of study selection for systematic review

---

*To tal number of included references is /three.tf/two.tf mapping to /five.tf unique trials

---
Effect of PSA screening on prostate-specific mortality

PSA screening may have little or no effect on prostate cancer-specific mortality based on five trials reporting this outcome (IRR 0.96 (0.85 to 1.08), I²=58%, low quality evidence; fig 4, table 2).10 23 31 33 This corresponds to zero fewer deaths from prostate cancer (95% CI 0 fewer to 0 more) per 1000 participants screened. We downgraded the quality of evidence for risk of bias as well as clinically important inconsistency.

Effect of PSA screening on incidence of prostate cancer

Based on data from four trials, screening may increase the detection of prostate cancer of any stage (IRR 1.23 (1.03 to 1.48), I²=99%, low quality evidence; fig 5, table 2).10 23 31 32 This corresponds to seven more diagnoses of prostate cancer (95% CI 1 more to 15 more) per 1000 men screened. We downgraded the quality of evidence for risk of bias as well as clinically important inconsistency.

PSA screening may increase the detection of localised (stage I and II) prostate cancer based on evidence from three randomised controlled trials (risk ratio (RR) 1.39 (1.09 to 1.79), I²=99%, low quality evidence; fig 6, table 2).10 31 32 This corresponds to seven more localised prostate cancers diagnosed (95% CI 2 more to 15 more) per 1000 participants screened. We downgraded the quality of evidence due to risk of bias and clinically important inconsistency. Based on the same body of evidence, PSA screening may modestly decrease the incidence of advanced prostate cancer (stage III and IV) (RR 0.85 (0.72 to 0.99), I²=87%, low quality evidence; fig 7, table 2). This corresponds to two fewer men diagnosed with advanced prostate cancer (95% CI 4 fewer to 0 fewer) per 1000 men screened. We downgraded the quality of evidence for risk of bias and clinically important inconsistency.

Subgroup analysis

We found no evidence of a subgroup effect according to age or screening frequency (appendix 3), but we did find a subgroup effect according to risk of bias (lower versus higher risk of bias).

Other outcomes

Only a single trial, the Finnish arm of the ERSPC trial15 provided data on quality of life when comparing PSA screening with no screening.15 31 This was based on a random sample of participants (n=1088) from both trial arms excluding men with a subsequent diagnosis of prostate cancer. Using the SF-6D instrument and a scale from 0 to 1.0 (with higher values indicating better quality of life), there was no difference between the two arms (mean difference 0.01 (95% CI 0.01 lower to 0.02 higher), low quality evidence; table 2). We downgraded the quality of evidence for risk of bias and indirectness.

Only the PLCO and CAP studies reported complications following biopsy.10 32 The PLCO study reported 75 complications from biopsies, including 29 infectious and 48 non-infectious adverse events. The CAP study reported three biopsy related complications (1 attributed to biopsy, 2 post-biopsy). Based on the Prostate Biopsy Effects cohort study, rates of biopsy-related complications ranged from 93% (haematospermia) to 1.4% (hospital readmissions, most commonly for sepsis)17: this corresponds to 94 men and one more man, respectively, per 1000 men screened. Graded by severity, 64.6% (95% CI 61.6% to 67.8%) experienced minor complications, 31.8% (28.8% to 35.1%) had moderate complications, and 1.4% (0.8% to 2.4%) had major complications. There were no biopsy related deaths. At six years after active monitoring, radical surgery, and local radiation, rates of urinary incontinence were 8%, 17%, and 4%, respectively, and rates of erectile dysfunction were 70%, 83%, and 73%, respectively (table 2).18 Comparing screened and unscreened men, we estimated there would be three more men per 1000 screened presenting with urinary incontinence and 25 more men per 1000 screened with erectile dysfunction (appendix 4).
Approximately two thirds of men with an elevated PSA level can expect a false positive test result, meaning they will not be diagnosed with prostate cancer. 

Approximately 15% of men with a PSA level $< 4$ ng/mL will harbour prostate cancer of any grade consistent with a false negative result. Clinically meaningful disease with a Gleason score $\geq 7$ can be expected in 2.3% of men with a PSA level $< 4$ ng/mL (table 2).

**Sensitivity analysis based on low risk of bias studies**

As planned, we performed an additional analysis using studies that were judged to be at lower risk of bias; this left only the ERSPC study.

**Effect of PSA screening on all-cause mortality**

Based on the ERSPC trial, PSA screening probably has no effect on all-cause mortality (IRR 1.0 (0.98 to 1.02), moderate quality evidence; table 3). This corresponds to zero fewer deaths of any cause (95% CI 3 fewer to 3 more) per 1000 participants screened. We downgraded the quality of evidence for risk of bias.

**Effect of PSA screening on prostate-specific mortality**

PSA screening probably has a small effect on prostate cancer-specific mortality (IRR 0.79 (0.69 to 0.91), moderate quality evidence; table 3). This corresponds to one fewer death from prostate cancer (95% CI 1 fewer to 0 fewer) per 1000 participants screened. We downgraded the quality of evidence for risk of bias. When using the ERSPC control event rate at 13 years or that of the Göteborg arm of the ERSPC trial at 18 years, this translates to one fewer (95% CI 1 fewer to 2 fewer) or three fewer (95% CI 1 fewer to 4 fewer) deaths from prostate cancer, respectively, per 1000 men screened.

**Effect of PSA screening on incidence of prostate cancer**

Screening probably increases the detection of prostate cancer of any stage (IRR 1.57 (1.51 to 1.62), moderate quality evidence; table 3). This corresponds to 18 more diagnoses of prostate cancer (95% CI 16 more to 20 more) per 1000 men screened. We downgraded the quality of evidence for risk of bias. Results of this analysis changed little when the Stockholm trial was excluded in a sensitivity analysis. Screening probably also increases the incidence of localised (stage I and II) prostate cancer (RR 1.75 (1.68 to 1.82)), corresponding to 14 more per 1000 (95% CI 13 more to 16 more), and probably decreases the incidence of advanced (stages III and IV) prostate cancer (RR 0.75 (0.69 to 0.82)), corresponding to three fewer per 1000 men screened (95% CI 4 fewer to 2 fewer).

**Other outcomes**

Since all available data for these outcomes were from the ERSPC trial the results of the sensitivity analysis were the same as that of the main analysis.

**Discussion**

**Statement of principal findings**

Based on moderate and low quality evidence, PSA screening seems to increase the detection of prostate cancer of any stage, increases the detection of stage I and II prostate cancer, and slightly decreases the detection of stage III and IV prostate cancer. At the same time, it probably modestly reduces prostate cancer specific mortality but has no effect on overall mortality. While findings from the ERSPC trial reflects a 21% relative risk reduction of prostate cancer-specific mortality (95% CI 9.0% to 31.0%), this corresponds to only one less death from prostate cancer (95% CI 1 fewer to 0 fewer) per 1000 men screened. Meanwhile,
PSA screening is associated with considerable biopsy-related and cancer treatment-related complications. We estimated that, for every 1000 men screened, approximately one, three, and 25 more men will be hospitalised for sepsis, require pads for urinary incontinence, and report erectile dysfunction, respectively.

**Strengths and weaknesses of the study**

We conducted this review based on an a priori protocol that defined a rigorous methodological approach based on the Cochrane Handbook and GRADE approach. Patient-centric outcomes and secondary analyses were informed by input from stakeholder representatives from the Rapid Recommendations guideline panel as well as a systematic review of the values and preferences of affected individuals. Our approach included a comprehensive search of multiple databases as well as other sources for relevant publications irrespective of language or publication status. While it is possible that we may have missed some secondary reports of the included trials, it seems unlikely that additional trials were missed.

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Summary of findings from pooled analysis of all eligible trials (n=5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outcome and timeframe</td>
<td>Study results and measurements</td>
</tr>
<tr>
<td>Mortality and cancer incidence at 10 years*</td>
<td></td>
</tr>
<tr>
<td>All-cause mortality</td>
<td>IR 0.99 (95% CI 0.98 to 1.01) based on data from 675 232 patients in 4 studies. Follow up 10-20 years</td>
</tr>
<tr>
<td>Prostate cancer mortality</td>
<td>IR 0.96 (0.85 to 1.08) based on data from 721 718 patients in 5 studies. Follow up 10-20 years</td>
</tr>
<tr>
<td>Incidence of prostate cancer (any stage)</td>
<td>IR 1.23 (1.03 to 1.48) based on data from 675 232 patients in 4 studies. Follow up 10-20 years</td>
</tr>
<tr>
<td>Incidence of localised prostate cancer (stages I and II)</td>
<td>IR 1.39 (1.09 to 1.79) based on data from 647 751 patients in 3 studies. Follow up 10-20 years</td>
</tr>
<tr>
<td>Incidence of advanced prostate cancer (stages III and IV)</td>
<td>IR 0.85 (0.72 to 0.99) based on data from 647 751 patients in 3 studies. Follow up 10-20 years</td>
</tr>
</tbody>
</table>

**Quality of life**

<table>
<thead>
<tr>
<th></th>
<th>Measured by SF-6D (scale 0-1, high better) based on data from 1088 patients in 1 study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quality of life</td>
<td>0.76 mean</td>
</tr>
</tbody>
</table>

**Complications and adverse effects**

<table>
<thead>
<tr>
<th></th>
<th>Most common complications were blood in semen (93%), blood in urine (66%), pain (44%), shivers (19%), and fever (18%). 1.4% (95% CI 0.8% to 2.4%) of men admitted to hospital (mostly for sepsis)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complication rates per biopsy</td>
<td>High (high quality and representative cohort study; intervention arm of CAP/ProtecT trial)</td>
</tr>
<tr>
<td>Biopsy related complications</td>
<td>Low (because of additional uncertainty due to estimating likelihood along the diagnostic pathway)</td>
</tr>
<tr>
<td>Complication rates by prostate cancer treatment modality</td>
<td>High (high quality prospective cohort studies of a 3-armed RCT)</td>
</tr>
<tr>
<td>Complications of subsequent prostate cancer treatment</td>
<td>Low (because of additional uncertainty due to estimating likelihood along the diagnostic pathway)</td>
</tr>
<tr>
<td>False positive screening results within 1 year</td>
<td>High (high quality prospective cohort from 5 arms of the ERSPC trial)</td>
</tr>
</tbody>
</table>

(continued)
The major limitation of this review stems from the included trials themselves. All trials had methodological limitations that lowered the confidence of their effect size estimates, as summarised in tables 2 and 3. Except for overall mortality, we found evidence of considerable inconsistency for each pooled analysis, prompting us to further downgrade the quality of the evidence. The key issue was considerable clinical and methodological heterogeneity across trials and within the arms of the ERSPC trial itself. The limited number of available trials precluded many planned secondary analyses of observed inconsistencies. Potential sources of heterogeneity include different screening intensity (one-time screen in CAP versus multiple rounds in other trials), different screening intervals (annual in PLCO versus every 2-4 years in different arms of the ERSPC trial, for example), different PSA biopsy thresholds (ranging from 2.5 to 4 ng/mL), as well as various degrees of adherence and control group contamination. In contrast to other systematic review authors,\textsuperscript{35, 36} we chose not to adjust for these inconsistencies as this would introduce additional uncertainty. Second, the reported harms evidence was not based on a comprehensive evaluation of the published literature, but instead used any available follow-up evidence from the intervention arms of the included trials and respective follow-up publications. This data was used to estimate the absolute number of biopsy and treatment related complications among men who underwent screening versus those who did not (see also infographic in linked Rapid Recommendation article\textsuperscript{12} and appendix 4).

The area of the square represents the weight given to the study in the meta-analysis. The pooled IRR was calculated by DerSimonian–Laird random effects model. The diamond represents the overall estimated effect and its 95% CI.

**Fig 3** | Forest plot showing the incidence rate ratio (IRR) for all-cause mortality for PSA screening v control groups. Horizontal bars denote 95% CIs. Studies are represented as squares centred on the point estimate of the result of each study. The area of the square represents the weight given to the study in the meta-analysis. The pooled IRR was calculated by DerSimonian–Laird random effects model. The diamond represents the overall estimated effect and its 95% CI.
Table 3 | Summary of findings from lower risk of bias data (sensitivity analysis based on ERSPC trial)

<table>
<thead>
<tr>
<th>Outcome and timeframe</th>
<th>Study results and measurements</th>
<th>Absolute effect estimates</th>
<th>Certainty in effect estimates (quality of evidence)</th>
<th>Plain text summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mortality and cancer incidence at 10 years*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All-cause mortality</td>
<td>IR 1.0 (95% CI 0.98 to 1.02) based on data from 162 243 patients in 1 study, Follow-up 13 years</td>
<td>129/1000 men</td>
<td>129/1000 men</td>
<td>0 (3 fewer to 3 more) Moderate (serious risk of bias)</td>
</tr>
<tr>
<td>Prostate cancer mortality</td>
<td>IR 0.79 (0.69 to 0.91) based on data from 162 243 patients in 1 study, Follow-up 13 years</td>
<td>3/1000 men</td>
<td>2/1000 men</td>
<td>1 fewer (1 fewer to 0) Moderate (serious risk of bias)</td>
</tr>
<tr>
<td>Incidence of prostate cancer (any stage)</td>
<td>IR 1.57 (1.51 to 1.62) based on data from 162 243 patients in 1 study, Follow-up 13 years</td>
<td>32/1000 men</td>
<td>50/1000 men</td>
<td>18 more (16 to 20 more) Moderate (serious risk of bias)</td>
</tr>
<tr>
<td>Incidence of localised prostate cancer (stages I and II)</td>
<td>IR 1.79 (1.68 to 1.82) based on data from 162 243 patients in 1 study, Follow-up 13 years</td>
<td>19/1000 men</td>
<td>33/1000 men</td>
<td>14 more (13 to 16 more) Moderate (serious risk of bias)</td>
</tr>
<tr>
<td>Incidence of advanced prostate cancer (stages III and IV)</td>
<td>IR 0.75 (0.69 to 0.82) based on data from 162 243 patients in 1 study, Follow-up 13 years</td>
<td>13/1000 men</td>
<td>10/1000 men</td>
<td>3 fewer (4 to 2 fewer) Moderate (serious risk of bias)</td>
</tr>
</tbody>
</table>

Quality of life

| Quality of life | Measured by SF-60 (scale 0-1, higher better) based on data from 1088 patients in 1 study | 0.76 mean | 0.75 mean | 0.01 lower (0.01 lower to 0.02 higher) Low (risk of bias and indirectness)§ | PSA screening may have little or no difference on quality of life |

Complications and adverse effects (identical to results in table 2)

IR=incidence ratio. RR=relative risk. PSA=prostate-specific antigen.

We provide the most up to date report on the best available evidence on screening for prostate cancer. While our group5 and others37 have previously published similar reviews, these did not include the most recent updates of existing screening trials nor the recently published CAP trial. The most relevant study for comparison is the systematic review and evidence report referenced by the recently published United States Public Services Task Force (USPSTF) recommendations on prostate cancer screening, which differs from our review in several important ways.9 38 First, the CAP trial—the largest and most recent trial of PSA screening to date—is missing from their evidence report. While the USPSTF guideline developers gave it formal consideration, the CAP trial was published after the USPSTF’s literature search cut-off date of 1 February 2018. Second, the authors excluded three older, “poor quality” trials based on methodological grounds. Reasons included low adherence to the screening

Fig 4 | Forest plot showing the incidence rate ratio (IRR) for prostate-specific mortality for PSA screening v control groups. Horizontal bars denote 95% CIs. Studies are represented as squares centred on the point estimate of the result of each study. The area of the square represents the weight given to the study in the meta-analysis. The pooled IRR was calculated by DerSimonian–Laird random effects model. The diamond represents the overall estimated effect and its 95% CI

Strengths and weaknesses in relation to other studies

We provide the most up to date report on the best available evidence on screening for prostate cancer. While our group5 and others37 have previously published similar reviews, these did not include the most recent updates of existing screening trials nor the recently published CAP trial. The most relevant study for comparison is the systematic review and evidence report referenced by the recently published United States Public Services Task Force (USPSTF) recommendations on prostate cancer screening, which differs from our review in several important ways.9 38 First, the CAP trial—the largest and most recent trial of PSA screening to date—is missing from their evidence report. While the USPSTF guideline developers gave it formal consideration, the CAP trial was published after the USPSTF’s literature search cut-off date of 1 February 2018. Second, the authors excluded three older, “poor quality” trials based on methodological grounds. Reasons included low adherence to the screening
intervention, uncertain levels of contamination in the control arm, and lack of blinding of outcome assessors. While we appreciate these methodological issues, their severity did not prompt us to disregard two of the three studies entirely as the USPSTF Task Force did. We excluded the Norköpping trial on a number of methodological limitations, including quasi-randomisation. Third, in contrast to the task force’s evidence report, we meta-analysed the entire body of evidence. This approach allowed us to evaluate for inconsistency between individual trials by conducting subgroup analyses (based on age and screening intervals) and a sensitivity analysis for risk of bias. Fourth, another major distinction lies in how we handled the control event rates. For the presentation of absolute effects in our summary of findings, we chose to draw this event rate from the CAP trial because it provides the most contemporary and broadly applicable estimation of what would occur in the absence of screening. These numbers were ascertained at a 10-year time horizon (while the effect estimates were pooled at the longest follow-up available in the included trials). This contributes to the lower absolute effect size estimates that we report compared with those reported by the USPSTF, but were deemed the most appropriate to inform the accompanying Rapid Recommendation panel deliberations.
Implications for clinicians and policy makers
This systematic review provides important information for an individual man’s decision making about prostate cancer screening. It supported the development of an evidence-based clinical practice guideline, as reflected in the accompanying Rapid Recommendation article.12 Our analysis indicates that PSA screening yields, at best, only a small benefit in prostate cancer specific mortality but does not reduce overall mortality. This small benefit should be weighed against the potential short term complications (biopsy related, false positive and false negative findings) and long term downstream effects (treatment related side effects, in particular related to urinary and sexual function). The latest results of the PIVOT and ProtecT trials should help guideline developers identify the subset of patients with clinically localised prostate cancer who will likely benefit from local, curative treatment while avoiding overtreatment.39-46 Although active surveillance is increasingly accepted to treat low risk prostate cancer, it is burdensome to patients due to the need for frequent follow-up visits, PSA testing, digital rectal examinations, prostate biopsies, and, recently, magnetic resonance imaging (MRI). The value of MRI in determining which prostate cancers appear clinically relevant versus indolent seems promising but remains uncertain in the context of PSA screening.41-43

AUTHOR AFFILIATIONS
1School of Public Health and Preventive Medicine, Monash University, Victoria, Australia
2National Clinicians Scholars Program, Yale University School of Medicine, and Veterans Affairs Connecticut Healthcare System, New Haven, Connecticut, USA
3Department of Urology, Yonsei University Wonju College of Medicine, Wonju, Korea
4Department of Urology, Chonnam National University Medical School, Gwangju, Korea
5Urology Section, Minneapolis VAMC and Department of Urology, University of Minnesota, Minneapolis, Minnesota, USA
6Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, ON, Canada
7Vellindre NHS Trust, Cardiff University Library Services, Vellindre Cancer Centre, Cardiff, Wales
8Division of General Internal Medicine and Division of Epidemiology, Department of Internal medicine, Rehabilitation and Geriatrics, University Hospitals of Geneva, Geneva, Switzerland

Contributors: DI, MD, and PD were involved in the conception and design of the review. DI, MD, AC, and PD developed the search strategy and performed study selection. DI, MD, JH, ECH, and PD extracted data from included studies. QZ, TA, and PD were involved in the data analysis. DI, MD, JH, ECH, QZ, TA, and PD were involved in the interpretation and discussion of results. All authors drafted the manuscript, contributed to the drafting of the review, and revised it critically for important intellectual content. All authors approved the final version of the article. All authors had access to all of the data in the study and can take responsibility for the integrity of the data and the accuracy of the data analysis. PD is guarantor.

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Competing interests: All authors have completed the Unified Competing Interest Form (available on request from the corresponding author) and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years, no other relationships or activities that could appear to have influenced the submitted work.

Ethical approval: Not required.

Data sharing: No additional data available.

Transparency: DI and PD affirm that the manuscript is an honest, accurate, and transparent account of the study being reported; no important aspects of the study have been omitted.

Patient involvement: As further described in the Methods section, three patient representatives helped inform the questions, outcomes and thresholds of clinical importance for this systematic review.

This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

18Donovan JL, Hamdy FC, Lane JA, et al, Protect Study Group*. Patient-reported outcomes after monitoring, surgery, or radiotherapy

Appendix 1: Search terms and strategies

Appendix 2: Characteristics of patients and trials included in the systematic review of the effects of PSA screening

Appendix 3: Subgroup analyses based on age, screening frequency, and risk of bias

Appendix 4: Modelling approach used to estimate the absolute number of biopsy and treatment related complications